

A Model Technological and Social Architecture for the Preservation of State Government Digital Information



Project manager Minnesota Historical Society (MHS)

Contact Robert Horton
State Archivist, Director of Library, Publications and Collections
345 Kellogg Blvd W
St. Paul, MN 55102
Voice: 651 259 3240, Fax: 651 296 9961
robert.horton@mnhs.org

Lead Partners California Digital Library (CDL), Patricia Cruse
National Conference of State Legislatures (NCSL), Jo Anne Bourquard
California Legislative Counsel, State Library and State Archives
Minnesota Revisor's Office, Legislative Reference Library and State Archives

State Partners Illinois (Connie Frankenfeld), Kansas (Matt Veatch), Mississippi (David Pilcher), Tennessee (Wayne Moore), Vermont (Gregory Sanford)

Problem Statement

Participants in the *Library of Congress Convening Workshops with the States* identified many compelling issues associated with state government digital preservation. One was how to prevent the loss of the large body of digital content already "at-risk." The 170 state and territorial librarians, archivists and records managers in attendance considered legislative records among those most "at-risk." Another challenge was the lack of capability to develop and use collective resources to address common needs.

Project description

This project addresses both challenges. It will emphasize working with a key audience, legislators and legislative staffs, and focusing on one of their key needs, providing enhanced access to multi-state collections. Previous work in Minnesota and California indicates that access is an effective catalyst for investment; by demonstrating immediate value to funding sources and to important constituencies, it becomes easier to justify and develop support for preservation. Correctly designed, the architecture for collecting, sharing and providing access to digital content will better position repositories to preserve that content.

To demonstrate that, the project will: 1) create a trustworthy technological and social architecture for capturing, managing and providing access to legislative content from California and Minnesota; 2) test the capacity of other states to adopt that model; 3) analyze the results with the legislative and government records communities; 4) define the thresholds for participation, to identify how different states could position themselves to implement similar programs; 5) use education and outreach to promote further collaboration and sustainability; and 6) develop compelling business cases for further investment.

The four state partners share a significant range of common ground, but have unique factors and opportunities to consider as well. Three, for example, are planning XML based bill drafting systems; one is designing its application with an Open Office front end. Each has a different potential, with disparate resources, possible partners and mandates, suggesting that implementation in each state will proceed from what is universal towards what is particular. Discussing and evaluating their situations will provide a larger perspective, anticipating the concerns of all the other states that will review this project.

Given that, the project deliverables will have to allow for modifications as appropriate in every state, identifying where content, expertise and responsibilities can be shared, as well as where more individual and appropriate solutions are necessary. In general, the conclusions should be:

- extensible, with options for multiple partners;
- cost-effective, with procedures simplified to the greatest extent possible;
- distributed, with the partners assuming appropriate responsibilities;
- useful, with a focus on material of clear value and return on investment;
- incremental, built on a modular approach;
- sustainable, based on regional collaborations and consortia; and
- practical, oriented towards specific communities with a demonstrable return on investment.

Estimated project cost

At a preliminary estimate, the project would cost \$600,000 over 24 months. This includes the cost of additional staff, expansion of technological capacity, meetings for partners and potential partners, education and outreach, hiring consultants and vendors. The project budget would meet all the costs of this effort, requesting only time and cooperation from the state partners and the NCSL. All the partners have indicated their willingness to commit to the project, contingent on the final plan and contract negotiated with the Library of Congress. The states, and any states added later, would look for additional funding from other sources – e.g., their own governments – to support the further development of capacity consequent to the project.

Project implementation

The project will build on current activities. The CDL has made a significant investment in digital preservation, which includes an operational Web Archiving Service (WAS). The core capabilities of the WAS are being extended to include the Web-at-Risk project, an archiving service that will enable memory institutions to select, capture, manage, and preserve Web-based government and political content. The MHS is now working with the Minnesota Revisor's Office and the California Legislative Counsel in the NHPRC funded E-Legislature project. This project appraised legislative records of all kinds and tested the transfer and preservation of content from XML based bill-drafting systems with different preservation architectures.¹ The NCSL is promoting the development of standards and means to help its members share information and expertise; it is currently working with the private sector to provide access to current legislative content and wants this project to support the preservation and provision of access to archival content.

The activities will include:

¹ The project website, with all current reports and products, is: <http://www.mnhs.org/elegislature>.

- **Access:** The MHS and CDL will develop and test web-based tools for providing access to content from California and Minnesota. The focus is on legislative content. The MHS will test the use of the federated search engine tool it is using, IDOL, an application from Autonomy, Inc. This will allow for testing of searches to both distributed and aggregated content. The selection of content will start with the appraisals done in the E-Legislature project, focusing on bills, acts, mandated reports and house and senate journals. The exact selection of content and the development of tools to enhance its value will be reviewed and validated in conjunction with meetings and focus groups facilitated with the NCSL, which represents the primary audiences for the use of legislative content; legislators and legislative staff are also a key constituency and source of funding. The content will be freely available online.
- **Evaluation:** MHS staff will work with four states to evaluate their capacity to work with their legislatures. The evaluation will examine both the solutions proposed by California and Minnesota, as well as the practices and opportunities in each state. The evaluation process will include: data consultant work, to appraise records and their digital manifestations; the review of workflow and data models; and a gap analysis mapping the appropriate next steps for each state.
- **Guidelines and standards:** Working with the states and the partners, MHS staff will translate what they have learned during the project into formal guidelines and standards. In conjunction with other products, such as the NDIPP/CTG tools for determining capacity, these will comprise a methodology for evaluating the preservation environment for legislative content in each state and the solutions appropriate to its resources and needs.
- **Preservation:** There are a variety of preservation models under development, two under the aegis of the NDIPP states initiative. Project staff will consult with a range of potential preservation partners to determine how those options will shape the standards and procedures pertinent to this effort. They will test the transfer of content to a number of the preservation options, to be determined in the course of the project. As part of the process, they will analyze the costs and benefits of each, to allow states to make the decisions about the appropriate solution for its needs. Minnesota and California will take formal stewardship of the content they collect in the project.
- **Business cases:** Any electronic records program has to attract support, especially given the need for collaborative solutions. The business case for managing legislative records will begin with their undoubted value and then put the access and preservation needs into the larger context of the related issues the states face. In Minnesota, for example, the debates over mandating open document formats or the CIO's plans to develop an enterprise-wide electronic document management system will have telling impact on this project. Each state archives has to understand the context in which it acts in order to make effective decisions; the MHS staff will develop a business case for Minnesota, along with a methodology for each state to develop its own.

- Promotion: The partners will promote the results of the project through a project website, maintained by the MHS, and through reports, presentations and meetings, as, for example, with the NCSL and NALIT.

Deliverables

Given the dynamic nature of technology, the various budgetary pressures on states, the quite disparate capacities of the state archives/libraries across the nation and the complexity of the issues at hand, no project will ever develop a single, permanent solution for the preservation of digital government records. But this project can demonstrate a viable, working solution that supports one of the most critical functions of government, as well as indicating how individual states can adopt and modify it.

The specific products and deliverables are:

- website, with enhanced search tools, providing access to archival digital content from California and Minnesota legislatures
- evaluations of capacity and opportunity in four states
- guidelines, standards, tools and methodology for other states to use to evaluate their own capacity to work with their legislatures
- evaluations of selected preservation options, with costs and benefits analyses
- business cases for investment in the preservation of legislative digital content
- promotion and dissemination - presentations and sessions at professional conferences and meetings, all reports and products online

This project aspires to position other states to turn their problems into their opportunities; it can lend them the knowledge and expertise they need to offer real, valued assistance to their potential partners in government. But it is important to note that this project can only position states to take advantage of these opportunities and to realize this potential; in other words, whatever we do, each state will have to take some additional steps on its own, but, with the support of this project, each state will know what those steps are, as well as what costs and benefits they represent, so that it can plan carefully and act confidently.